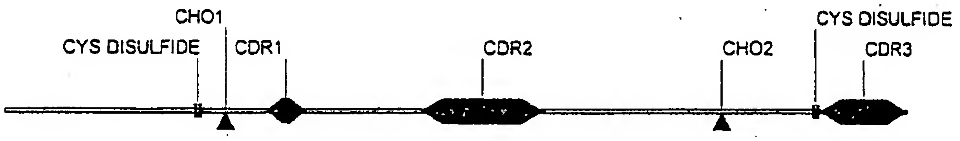
Figure 1A. Alignment of Thy-1 and 8E5 VH

```
                           1                                                 50
EcFv-1.15* 8E5 VH    (1)   --------------QLQQS--G--AELVKPGASVKMSCKASGYTFTNYWM
       Thy1 human    (1)   MNLAISIALLLTVLQVSRGQKVTSLTACLVDQSLRLDCRHENTSSSPIQY
        Consensus    (1)                 QL                     SLKL CK     S S
                           51                                               100
EcFv-1.15* 8E5 VH   (33)   HWVKQRPGQGLEWIGTIDPADSYTSYNQNF----------------KDKA
       Thy1 human   (51)   EESLTRETKKHVLFGTVGVPEHTYRSRTNFTSKYHMKVLYLSAFTSKDEG
        Consensus   (51)    F    R        GTI    D        NF                KD A
                           101                                              150
EcFv-1.15* 8E5 VH   (67)   TLTVDKPSSTAYMQLSSLTFGDSAVYFCAREGVYYRYYFDYWGHGTTLTV
       Thy1 human  (101)   TYTCALHHSGHSPPISSQNVTVLRDKLVKCEGISLLAQNTSWLLLLLLSL
        Consensus  (101)   T T     S     ISS             EGI         W     LSL
                           151       161
EcFv-1.15* 8E5 VH  (117)   SSAKTTPKL--
       Thy1 human  (151)   SLLQATDFMSL
        Consensus  (151)   S    T  L
```

Figure 1B. Design of a single Ig domain CDR binding polypeptide based on the Thy-1 structure.

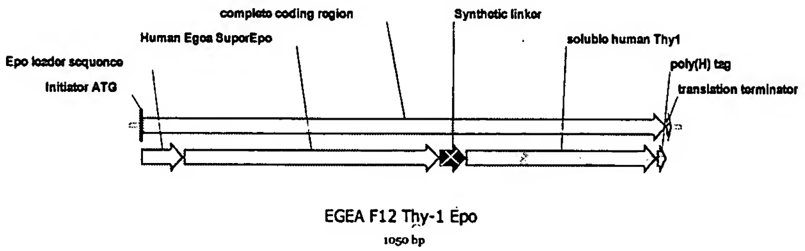
Qvsrgqkvtsltaclvdqslrldcrhentsssnywm
Hfsltretkkhvlfgtidpadsytsynqnfkdegtytc
Alhhsghsppissqnvtvlrdklvkcegvyyyryyfdy

Figure 1C. Diagram of a single Ig domain CDR carrier based on the Thy-1 structure.



Thy1/8E4 VH synthetic CDR binding polypeptide
111 aa

A

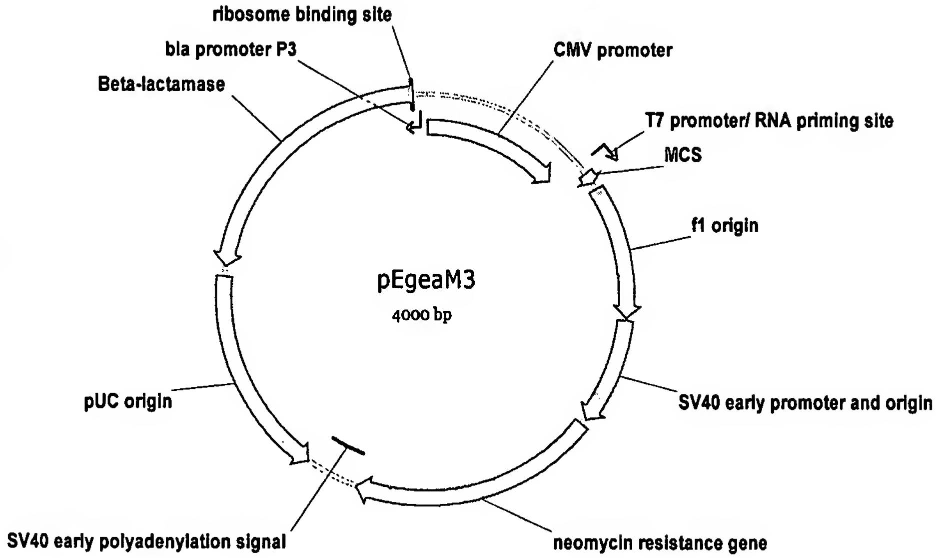


B

```
                              MetGlyValHis GluCysPro AlaTrpLeu TrpLeuLeuLeu SerLeuLeu SerLeuPro LeuGlyLeuPro ValLeuGly·
   1    GATTGGCGAA GCTTGGAGGA ATGGGCGTGC ACGAGTGCCC CGCCTGGCTG TGGCTGCTGC TGAGCCTGCT GAGCCTGCCC CTGGGCCTGC CCGTGCTGGG
        CTAACCGCTT CGAACCTCCT TACCCGCACG TGCTCACGGG GCGGACCGAC ACCGACGACG ACTCGGACGA CTCGGACGGG GACCCGGACG GGCACGACCC
        ·GlyProPro ArgLeuIleCys AspSerArg ValLeuGlu ArgHisLeuLeu GluAlaLys GluAlaGlu SerIleThrThr GlyCysVal GluAspCys
 101    CGCCCCCCCC CGGCTGATCT GCGACAGCCG GGTGCTGGAG CGGCACCTGC TGGAGGCCAA GGAGGCCGAG AGCATCACCA CCGGCTGCGT GGAGGACTGC
        GCGGGGGGGG GCCGACTAGA CGCTGTCGGC CCACGACCTC GCCGTGGACG ACCTCCGGTT CCTCCGGCTC TCGTAGTGGT GGCCGACGCA CCTCCTGACG
        SerLeuAsnGlu AsnIleThr ValProAsp SerLysValAsn PheTyrAla TrpLysArg MetGluValGly GlnGlnAla ValGluVal TrpGlnGlyLeu·
 201    AGCCTGAACG AGAACATCAC CGTGCCCGAC AGCAAGGTGA ACTTCTACGC CTGGAAGCGG ATGGAGGTGG GCCAGCAGGC CGTGGAGGTG TGGCAGGGCC
        TCGGACTTGC TCTTGTAGTG GCACGGGCTG TCGTTCCACT TGAAGATGCG GACCTTCGCC TACCTCCACC CGGTCGTCCG GCACCTCCAC ACCGTCCCGG
        ·LAlaLeuLeu SerGluAla ValLeuArgGly GlnAlaLeu LeuValIle SerSerGlnPro TrpGluPro LeuGlnLeu HisValAspLys AlaValSer·
 301    TGGCCCTGCT GAGCGAGGCC GTGCTGCGGG GCCAGGCCCT GCTGGTGATC AGCAGCCAGC CCTGGGAGCC CCTGCAGCTG CACGTGGACA AGGCCGTGAG
        ACCGGGACGA CTCGCTCCGG CACGACGCCC CGGTCCGGGA CGACCACTAG TCGTCGGTCG GGACCCTCGG GGACGTCGAC GTGCACCTGT TCCGGCACTC
        ·GlyLeuArg SerLeuThrThr LeuLeuArg AlaLeuGly AlaGlnLysGlu AlaIleSer ProProAsp AlaAlaSerAla AlaProLeu ArgThrIle
 401    CGGCCTGCGG AGCCTGACCA CCCTGCTGCG GGCCCTGGGC GCCCAGAAGG AGGCCATCAG CCCCCCCGAC GCCGCCAGCG CCGCCCCCCT GCGGACCATC
        GCCGGACGCC TCGGACTGGT GGGACGACGC CCGGGACCCG CGGGTCTTCC TCCGGTAGTC GGGGGGGCTG CGGCGGTCGC GGCGGGGGGA CGCCTGGTAG
        ThrAlaAspThr PheArgLys LeuPheArg ValTyrProAsn PheLeuArg GlyLysLeu LysPheTyrThr GlyGluAla CysArgGly GlyGlyGlyGly·
 501    ACCGCCGACA CCTTCCGGAA GCTGTTCCGG GTGTACCCCA ACTTCCTGCG GGGCAAGCTG AAGTTCTACA CCGGCGAGGC CTGCCGGGGC GGCGGCGGCG
        TGGCGGCTGT GGAAGGCCTT CGACAAGGCC CACATGGGGT TGAAGGACGC CCCGTTCGAC TTCAAGATGT GGCCGCTCCG GACGGCCCCG CCGCCGCCGC
        ·GSerGlyGly GlyGlyGlu PheGlyGlyGly GlySerGln LysValThr SerLeuThrAla CysLeuVal AspGlnSer LeuArgLeuAsp CysArgHis·
 601    GCAGCGGCGG CGGCGGCGAG TTCGGCGGCG GCGGCAGCCA GAAGGTGACC AGCCTGACCG CCTGCCTGGT GGACCAGAGC CTGCGGCTGG ACTGCCGGCA
        CGTCGCCGCC GCCGCCGCTC AAGCCGCCGC CGCCGTCGGT CTTCCACTGG TCGGACTGGC GGACGGACCA CCTGGTCTCG GACGCCGACC TGACGGCCGT
        ·GluAsnThr SerSerSerPro IleGlnTyr GluPheSer LeuThrArgGlu ThrLysLys HisValLeu PheGlyThrVal GlyValPro GluHisThr
 701    CGAGAACACC AGCAGCAGCC CCATCCAGTA CGAGTTCAGC CTGACCCGGG AGACCAAGAA GCACGTGCTG TTCGGCACCG TGGGCGTGCC CGAGCACACC
        GCTCTTGTGG TCGTCGTCGG GGTAGGTCAT GCTCAAGTCG GACTGGGCCC TCTGGTTCTT CGTGCACGAC AAGCCGTGGC ACCCGCACGG GCTCGTGTGG
        TyrArgSerArg ThrAsnPhe ThrSerLys TyrHisMetLys ValLeuTyr LeuSerAla PheThrSerLys AspGluGly ThrTyrThr CysAlaLeuHis·
 801    TACCGGAGCC GGACCAACTT CACCAGCAAG TACCACATGA AGGTGCTGTA CCTGAGCGCC TTCACCAGCA AGGACGAGGG CACCTACACC TGCGCCCTGC
        ATGGCCTCGG CCTGGTTGAA GTGGTCGTTC ATGGTGTACT TCCACGACAT GGACTCGCGG AAGTGGTCGT TCCTGCTCCC GTGGATGTGG ACGCGGGACG
        ·HHisSerGly HisSerPro ProIleSerSer GlnAsnVal ThrValLeu ArgAspLysLeu ValLysCys GluGlyIle SerLeuLeuAla GlnAsnThr·
 901    ACCACAGCGG CCACAGCCCC CCCATCAGCA GCCAGAACGT GACCGTGCTG CGGGACAAGC TGGTGAAGTG CGAGGGCATC AGCCTGCTGG CCCAGAACAC
        TGGTGTCGCC GGTGTCGGGG GGGTAGTCGT CGGTCTTGCA CTGGCACGAC GCCCTGTTCG ACCACTTCAC GCTCCCGTAG TCGGACGACC GGGTCTTGTG
        ·SerHisHis HisHisHisHis
1001    CAGCCACCAC CACCACCACC ACTGATGATA AGATCGGATC CTAGGCTTCC
        GTCGGTGGTG GTGGTGGTGG TGACTACTAT TCTAGCCTAG GATCCGAAGG
```

FIGURE 3

A



B

pEgea M3

```
  1   GATTATTCTA GACCCGCGTT ACATAACTTA CGGTAAATGG CCCGCCTGGC TGACCGCCCA ACGACCCCCG CCCATTGACG TCAATAATGA CGTATGTTCC
      CTAATAAGAT CTGGGCGCAA TGTATTGAAT GCCATTTACC GGGCGGACCG ACTGGCGGGT TGCTGGGGGC GGGTAACTGC AGTTATTACT GCATACAAGG
101   CATAGTAACG CCAATAGGGA CTTTCCATTG ACGTCAATGG GTGGAGTATT TACGGTAAAC TGCCCACTTG GCAGTACATC AAGTGTATCA TATGCCAAGT
      GTATCATTGC GGTTATCCCT GAAAGGTAAC TGCAGTTACC CACCTCATAA ATGCCATTTG ACGGGTGAAC CGTCATGTAG TTCACATAGT ATACGGTTCA
201   ACGCCCCCTA TTGACGTCAA TGACGGTAAA TGGCCCGCCT GGCATTATGC CCAGTACATG ACCTTATGGG ACTTTCCTAC TTGGCAGTAC ATCTACGTAT
      TGCGGGGGAT AACTGCAGTT ACTGCCATTT ACCGGGCGGA CCGTAATACG GGTCATGTAC TGGAATACCC TGAAAGGATG AACCGTCATG TAGATGCATA
301   TAGTCATCGC TATTACCATG GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTTGACTC ACGGGGATTT CCAAGTCTCC ACCCCATTGA
      ATCAGTAGCG ATAATGGTAC CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACTGAG TGCCCCTAAA GGTTCAGAGG TGGGGTAACT
401   CGTCAATGGG AGTTTGTTTT GGCACCAAAA TCAACGGGAC TTTCCAAAAT GTCGTAACAA CTCCGCCCCA TTGACGCAAA TGGGCGGTAG GCGTGTACGG
      GCAGTTACCC TCAAACAAAA CCGTGGTTTT AGTTGCCCTG AAAGGTTTTA CAGCATTGTT GAGGCGGGGT AACTGCGTTT ACCCGCCATC CGCACATGCC
501   TGGGAGGTCT ATATAAGCAG AGCTCTCTGG CTAACTAGAA TCGAAATTAA TACGACTCAC TATAGGGAGA CCCAAGCTGG CTAGCGTTTA AACTTAAGCT
      ACCCTCCAGA TATATTCGTC TCGAGAGACC GATTGATCTT AGCTTTAATT ATGCTGAGTG ATATCCCTCT GGGTTCGACC GATCGCAAAT TTGAATTCGA
601   TGGTACCGAG CTCGGATCCA CTCTAGGGGG TATCCCCACG CGCCCTGTAG CGGCGCATTA AGCGCGGCGG GTGTGGTGGT TACGCGCAGC GTGACCGCTA
      ACCATGGCTC GAGCCTAGGT GAGATCCCCC ATAGGGGTGC GCGGGACATC GCCGCGTAAT TCGCGCCGCC CACACCACCA ATGCGCGTCG CACTGGCGAT
701   CACTTGCCAG CGCCCTAGCG CCCGCTCCTT TCGCTTTCTT CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC GGGGGCTCCC
      GTGAACGGTC GCGGGATCGC GGGCGAGGAA AGCGAAAGAA GGGAAGGAAA GAGCGGTGCA AGCGGCCGAA AGGGGCAGTT CGAGATTTAG CCCCCGAGGG
801   TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAAACTTG ATTAGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT
      AAATCCCAAG GCTAAATCAC GAAATGCCGT GGAGCTGGGG TTTTTTGAAC TAATCCCACT ACCAAGTGCA TCACCCGGTA GCGGGACTAT CTGCCAAAAA
901   CGCCCTTTGA CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTTCCA AACTGGAACA ACACTCAACC CTATCTCGGT CTATTCTTTT GATTTATAAG
      GCGGGAAACT GCAACCTCAG GTGCAAGAAA TTATCACCTG AGAACAAGGT TTGACCTTGT TGTGAGTTGG GATAGAGCCA GATAAGAAAA CTAAATATTC
```

FIGURE 5

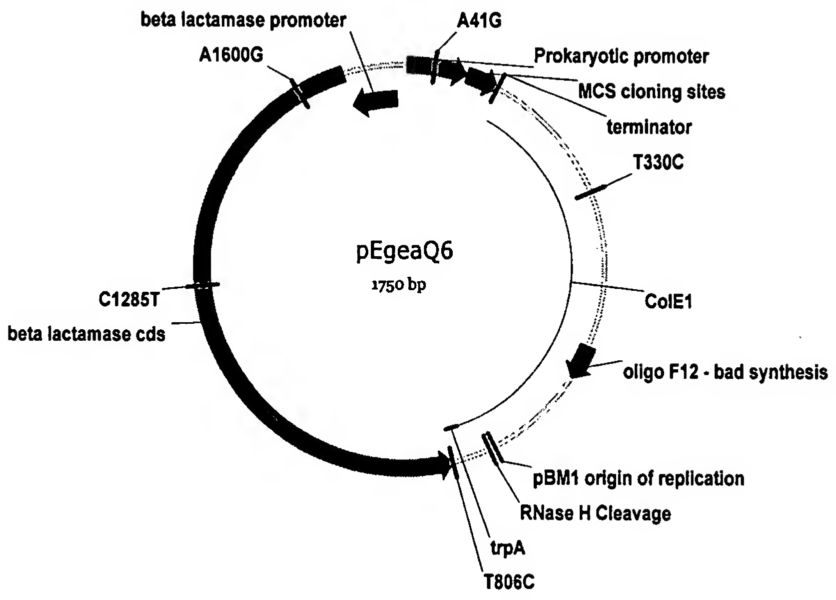
```
1001     GGATTTTGCC GATTTCGGCC TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTAATT CTGTGGAATG TGTGTCAGTT AGGGTGTGGA
         CCTAAAACGG CTAAAGCCGG ATAACCAATT TTTTACTCGA CTAAATTGTT TTTAAATTGC GCTTAATTAA GACACCTTAC ACACAGTCAA TCCCACACCT
1101     AAGTCCCCAG GCTCCCCAGC AGGCAGAAGT ATGCAAAGCA TGCATCTCAA TTAGTCAGCA ACCAGGTGTG GAAAGTCCCC AGGCTCCCCA GCAGGCAGAA
         TTCAGGGGTC CGAGGGGTCG TCCGTCTTCA TACGTTTCGT ACGTAGAGTT AATCAGTCGT TGGTCCACAC CTTTCAGGGG TCCGAGGGGT CGTCCGTCTT
1201     GTATGCAAAG CATGCATCTC AATTAGTCAG CAACCATAGT CCCGCCCCTA ACTCCGCCCA TCCCGCCCCT AACTCCGCCC AGTTCCGCCC ATTCTCCGCC
         CATACGTTTC GTACGTAGAG TTAATCAGTC GTTGGTATCA GGGCGGGGAT TGAGGCGGGT AGGGCGGGGA TTGAGGCGGG TCAAGGCGGG TAAGAGGCGG
1301     CCATGGCTGA CTAATTTTTT TTATTTATGC AGAGGCCGAG GCCGCCTCTG CCTCTGAGCT ATTCCAGAAG TAGTGAGGAG GCTTTTTTGG AGGCCTAGGC
         GGTACCGACT GATTAAAAAA AATAAATACG TCTCCGGCTC CGGCGGAGAC GGAGACTCGA TAAGGTCTTC ATCACTCCTC CGAAAAAACC TCCGGATCCG
1401     TTTTGCAAAA AGCTCGAGGA TCGTTTCGCA TGATTGAACA AGATGGATTG CACGCAGGTT CTCCGGCCGC TTGGGTGGAG AGGCTATTCG GCTATGACTG
         AAAACGTTTT TCGAGCTCCT AGCAAAGCGT ACTAACTTGT TCTACCTAAC GTGCGTCCAA GAGGCCGGCG AACCCACCTC TCCGATAAGC CGATACTGAC
1501     GGCACAACAG ACAATCGGCT GCTCTGATGC CGCCGTGTTC CGGCTGTCAG CGCAGGGGCG CCCGGTTCTT TTTGTCAAGA CCGACCTGTC CGGTGCCCTG
         CCGTGTTGTC TGTTAGCCGA CGAGACTACG GCGGCACAAG GCCGACAGTC GCGTCCCCGC GGGCCAAGAA AAACAGTTCT GGCTGGACAG GCCACGGGAC
1601     AATGAACTGC AGGACGAGGC AGCGCGGCTA TCGTGGCTGG CCACGACGGG CGTTCCTTGC GCAGCTGTGC TCGACGTTGT CACTGAAGCG GGAAGGGACT
         TTACTTGACG TCCTGCTCCG TCGCGCCGAT AGCACCGACC GGTGCTGCCC GCAAGGAACG CGTCGACACG AGCTGCAACA GTGACTTCGC CCTTCCCTGA
1701     GGCTGCTATT GGGCGAAGTG CCGGGGCAGG ATCTCCTGTC ATCTCACCTT GCTCCTGCCG AGAAAGTATC CATCATGGCT GATGCAATGC GGCGGCTGCA
         CCGACGATAA CCCGCTTCAC GGCCCCGTCC TAGAGGACAG TAGAGTGGAA CGAGGACGGC TCTTTCATAG GTAGTACCGA CTACGTTACG CCGCCGACGT
1801     TACGCTTGAT CCGGCTACCT GCCCATTCGA CCACCAAGCG AAACATCGCA TCGAGCGAGC ACGTACTCGG ATGGAAGCCG GTCTTGTCGA TCAGGATGAT
         ATGCGAACTA GGCCGATGGA CGGGTAAGCT GGTGGTTCGC TTTGTAGCGT AGCTCGCTCG TGCATGAGCC TACCTTCGGC CAGAACAGCT AGTCCTACTA
1901     CTGGACGAAG AGCATCAGGG GCTCGCGCCA GCCGAACTGT TCGCCAGGCT CAAGGCGCGC ATGCCCGACG GCGAGGATCT CGTCGTGACC CATGGCGATG
         GACCTGCTTC TCGTAGTCCC CGAGCGCGGT CGGCTTGACA AGCGGTCCGA GTTCCGCGCG TACGGGCTGC CGCTCCTAGA GCAGCACTGG GTACCGCTAC
2001     CCTGCTTGCC GAATATCATG GTGGAAAATG GCCGCTTTTC TGGATTCATC GACTGTGGCC GGCTGGGTGT GGCGGACCGC TATCAGGACA TAGCGTTGGC
         GGACGAACGG CTTATAGTAC CACCTTTTAC CGGCGAAAAG ACCTAAGTAG CTGACACCGG CCGACCCACA CCGCCTGGCG ATAGTCCTGT ATCGCAACCG
2101     TACCCGTGAT ATTGCTGAAG AGCTTGGCGG CGAATGGGCT GACCGCTTCC TCGTGCTTTA CGGTATCGCC GCTCCCGATT CGCAGCGCAT CGCCTTCTAT
         ATGGGCACTA TAACGACTTC TCGAACCGCC GCTTACCCGA CTGGCGAAGG AGCACGAAAT GCCATAGCGG CGAGGGCTAA GCGTCGCGTA GCGGAAGATA
2201     CGCCTTCTTG ACGAGTTCTT CTGAGCGGGA CGCACCCCAA CTTGTTTATT GCAGCTTATA ATGGTTACAA ATAAAGCAAT AGCATCACAA ATTTCACAAA
         GCGGAAGAAC TGCTCAAGAA GACTCGCCCT GCGTGGGGTT GAACAAATAA CGTCGAATAT TACCAATGTT TATTTCGTTA TCGTAGTGTT TAAAGTGTTT
2301     TAAAGCATTT TTTTCACTGC ATTCTAGTTG TGGTTTGTCC AAACTCATCA ATGTATCTTA TCATGTCTGT ATACCGTCGA CCTCTAGCTA ATGTGAGCAA
         ATTTCGTAAA AAAAGTGACG TAAGATCAAC ACCAAACAGG TTTGAGTAGT TACATAGAAT AGTACAGACA TATGGCAGCT GGAGATCGAT TACACTCGTT
2401     AAGGCCAGCA AAAGGCCAGG AACCGTAAAA AGGCCGCGTT GCTGGCGTTT TTCCATAGGC TCCGCCCCCC TGACGAGCAT CACAAAAATC GACGCTCAAG
         TTCCGGTCGT TTTCCGGTCC TTGGCATTTT TCCGGCGCAA CGACCGCAAA AAGGTATCCG AGGCGGGGGG ACTGCTCGTA GTGTTTTTAG CTGCGAGTTC
2501     TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC TCTCCTGTTC CGACCCTGCC GCTTACCGGA
         AGTCTCCACC GCTTTGGGCT GTCCTGATAT TTCTATGGTC CGCAAAGGGG GACCTTCGAG GGAGCACGCG AGAGGACAAG GCTGGGACGG CGAATGGCCT
2601     TACCTGTCCG CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCATAGCTC ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT
         ATGGACAGGC GGAAAGAGGG AAGCCCTTCG CACCGCGAAA GAGTATCGAG TGCGACATCC ATAGAGTCAA GCCACATCCA GCAAGCGAGG TTCGACCCGA
2701     GTGTGCACGA ACCCCCCGTT CAGCCCGACC GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC CACTGGCAGC
         CACACGTGCT TGGGGGGCAA GTCGGGCTGG CGACGCGGAA TAGGCCATTG ATAGCAGAAC TCAGGTTGGG CCATTCTGTG CTGAATAGCG GTGACCGTCG
2801     AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG AACAGTATTT
         TCGGTGACCA TTGTCCTAAT CGTCTCGCTC CATACATCCG CCACGATGTC TCAAGAACTT CACCACCGGA TTGATGCCGA TGTGATCTTC TTGTCATAAA
2901     GGTATCTGCG CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA CCACCGCTGG TAGCGGTTTT TTTGTTTGCA
         CCATAGACGC GAGACGACTT CGGTCAATGG AAGCCTTTTT CTCAACCATC GAGAACTAGG CCGTTTGTTT GGTGGCGACC ATCGCCAAAA AAACAAACGT
3001     AGCAGCAGAT TACGCGCAGA AAAAAAGGAT CTCAAGAAGA TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAAACCAGT TACCAATGCT
         TCGTCGTCTA ATGCGCGTCT TTTTTTCCTA GAGTTCTTCT AGGAAACTAG AAAAGATGCC CCAGACTGCG AGTCACCTTG CTTTTGGTCA ATGGTTACGA
3101     TAATCAGTGA GGCACCTATC TCAGCGATCT GTCTATTTCG TTCATCCATA GTTGCCTGAC TCCCCGTCGT GTAGATAACT ACGATACGGG AGGGCTTACC
         ATTAGTCACT CCGTGGATAG AGTCGCTAGA CAGATAAAGC AAGTAGGTAT CAACGGACTG AGGGGCAGCA CATCTATTGA TGCTATGCCC TCCCGAATGG
3201     ATCTGGCCCC AGTGCTGCAA TGATACCGCG AGACCCACGC TCACCGGCTC CAGATTTATC AGCAATAAAC CAGCCAGCCG GAAGGGCCGA GCGCAGAAGT
         TAGACCGGGG TCACGACGTT ACTATGGCGC TCTGGGTGCG AGTGGCCGAG GTCTAAATAG TCGTTATTTG GTCGGTCGGC CTTCCCGGCT CGCGTCTTCA
3301     GGTCCTGCAA CTTTATCCGC CTCCATCCAG TCTATTAATT GTTGCCGGGA AGCTAGAGTA AGTAGTTCGC CAGTTAATAG TTTGCGCAAC GTTGTTGCCA
```

FIGURE 5B CONT.

```
        CCAGGACGTT GAAATAGGCG GAGGTAGGTC AGATAATTAA CAACGGCCCT TCGATCTCAT TCATCAAGCG GTCAATTATC AAACGCGTTG CAACAACGGT
3401    TTGCTACAGG CATCGTGGTG TCACGCTCGT CGTTTGGTAT GGCTTCATTC AGCTCCGGTT CCCAACGATC AAGGCGAGTT ACATGATCCC CCATGTTGTG
        AACGATGTCC GTAGCACCAC AGTGCGAGCA GCAAACCATA CCGAAGTAAG TCGAGGCCAA GGGTTGCTAG TTCCGCTCAA TGTACTAGGG GGTACAACAC
3501    CAAAAAAGCG GTTAGCTCCT TCGGTCCTCC GATCGTTGTC AGAAGTAAGT TGGCCGCAGT GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT
        GTTTTTTCGC CAATCGAGGA AGCCAGGAGG CTAGCAACAG TCTTCATTCA ACCGGCGTCA CAATAGTGAG TACCAATACC GTCGTGACGT ATTAAGAGAA
3601    ACTGTCATGC CATCCGTAAG ATGCTTTTCT GTGACTGGTG AGTACTCAAC CAAGTCATTC TGAGAATAGT GTATGCGGCG ACCGAGTTGC TCTTGCCCGG
        TGACAGTACG GTAGGCATTC TACGAAAAGA CACTGACCAC TCATGAGTTG GTTCAGTAAG ACTCTTATCA CATACGCCGC TGGCTCAACG AGAACGGGCC
3701    CGTCAATACG GGATAATACC GCGCCACATA GCAGAACTTT AAAAGTGCTC ATCATTGGAA AACGTTCTTC GGGGCGAAAA CTCTCAAGGA TCTTACCGCT
        GCAGTTATGC CCTATTATGG CGCGGTGTAT CGTCTTGAAA TTTTCACGAG TAGTAACCTT TTGCAAGAAG CCCCGCTTTT GAGAGTTCCT AGAATGGCGA
3801    GTTGAGATCC AGTTCGATGT AACCCACTCG TGCACCCAAC TGATCTTCAG CATCTTTTAC TTTCACCAGC GTTTCTGGGT GAGCAAAAAC AGGAAGGCAA
        CAACTCTAGG TCAAGCTACA TTGGGTGAGC ACGTGGGTTG ACTAGAAGTC GTAGAAAATG AAAGTGGTCG CAAAGACCCA CTCGTTTTTG TCCTTCCGTT
3901    AATGCCGCAA AAAAGGGAAT AAGGGCGACA CGGAAATGTT GAATACTCAT ACTCTTCCTT TTTCAATATT ATTGAAGCAT TTATCTAGAG GTTATTGTCT
        TTACGGCGTT TTTTCCCTTA TTCCCGCTGT GCCTTTACAA CTTATGAGTA TGAGAAGGAA AAAGTTATAA TAACTTCGTA AATAGATCTC CAATAACAGA
```

FIGURE 5B CONT.

A



B

```
  1   AGTGCTCTAG ACCTGTTGAC AATTAATCAT CGGCTCGTAT AATGTGTGGA ATTGTGAGCG GATAACAATT TCACACAGGA AACAGGATCG ATCGAATTCG
      TCACGAGATC TGGACAACTG TTAATTAGTA GCCGAGCATA TTACACACCT TAACACTCGC CTATTGTTAA AGTGTGTCCT TTGTCCTAGC TAGCTTAAGC
101   GATCCAAGCT TGAGCTCGAG CCATGGCCCG GGTGAATAAT TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC
      CTAGGTTCGA ACTCGAGCTC GGTACCGGGC CCACTTATTA ATCTTTTCTA GTTTCCTAGA AGAACTCTAG GAAAAAAAGA CGCGCATTAG ACGACGAACG
201   AAACAAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC TTTTTCCGAA GGTAACTGGC TTCAGCAGAG CGCAGATACC
      TTTGTTTTTT TGGTGGCGAT GGTCGCCACC AAACAAACGG CCTAGTTCTC GATGGTTGAG AAAAAGGCTT CCATTGACCG AAGTCGTCTC GCGTCTATGG
301   AAATACTGTT CTTCTAGTGT AGCCGTAGTT AGGCCACCAC TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT
      TTTATGACAA GAAGATCACA TCGGCATCAA TCCGGTGGTG AAGTTCTTGA GACATCGTGG CGGATGTATG GAGCGAGACG ATTAGGACAA TGGTCACCGA
401   GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTCGGGCTG AACGGGGGGT TCGTGCACAC
      CGACGGTCAC CGCTATTCAG CACAGAATGG CCCAACCTGA GTTCTGCTAT CAATGGCCTA TTCCGCGTCG CCAGCCCGAC TTGCCCCCCA AGCACGTGTG
501   AGCCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA GGGACAAAGG CGGACAGGTA
      TCGGGTCGAA CCTCGCTTGC TGGATGTGGC TTGACTCTAT GGATGTCGCA CTCGATACTC TTTCGCGGTG CGAAGGGCTT CCCTGTTTCC GCCTGTCCAT
601   TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG TCGGGTTTCG CCACCTCTGA
      AGGCCATTCG CCGTCCCAGC CTTGTCCTCT CGCGTGCTCC CTCGAAGGTC CCCCTTTGCG GACCATAGAA ATATCAGGAC AGCCCAAAGC GGTGGAGACT
701   CTTGAGCGTC GATTTTTGTG ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC AACGCGGCCT TTTTACGGTT CCTGCCCGCT CATTAGGCGG
      GAACTCGCAG CTAAAAACAC TACGAGCAGT CCCCCCGCCT CGGATACCTT TTTGCGGTCG TTGCGCCGGA AAAATGCCAA GGACGGGCGA GTAATCCGCC
801   GCTATTACCA ATGCTTAATC AGTGAGGCAC CTATCTCAGC GATCTGTCTA TTTCGTTCAT CCATAGCTGC CTGACTCCCC GTCGTGTAGA TAACTACGAT
      CGATAATGGT TACGAATTAG TCACTCCGTG GATAGAGTCG CTAGACAGAT AAAGCAAGTA GGTATCGACG GACTGAGGGG CAGCACATCT ATTGATGCTA
901   ACGGGAGGGC TTACCATCTG GCCCCAGTGC TGCAATGATA CCGCGAGACC CACGCTCACC GGCTCCAGAT TTATCAGCAA TAAACCAGCC AGCCGGAAGG
      TGCCCTCCCG AATGGTAGAC CGGGGTCACG ACGTTACTAT GGCGCTCTGG GTGCGAGTGG CCGAGGTCTA AATAGTCGTT ATTTGGTCGG TCGGCCTTCC
```

FIGURE 6

```
1001     GCCGAGCGCA GAAGTGGTCC TGCAACTTTA TCCGCCTCCA TCCAGTCTAT TAATTGTTGC CGGGAAGCTA GAGTAAGTAG TTCGCCAGTT AATAGTTTGC
         CGGCTCGCGT CTTCACCAGG ACGTTGAAAT AGGCGGAGGT AGGTCAGATA ATTAACAACG GCCCTTCGAT CTCATTCATC AAGCGGTCAA TTATCAAACG
1101     GCAACGTTGT TGCCATTGCT ACAGGCATCG TGGTGTCACG CTCGTCGTTT GGTATGGCTT CATTCAGCTC CGGTTCCCAA CGATCAAGGC GAGTTACATG
         CGTTGCAACA ACGGTAACGA TGTCCGTAGC ACCACAGTGC GAGCAGCAAA CCATACCGAA GTAAGTCGAG GCCAAGGGTT GCTAGTTCCG CTCAATGTAC
1201     ATCCCCCATG TTGTGCAAAA AAGCGGTTAG CTCCTTCGGT CCTCCGATCG TTGTCAGAAG TAAGTTGGCC GCAGTGTTAT CACTCATGGT TATGGCAGCA
         TAGGGGGTAC AACACGTTTT TTCGCCAATC GAGGAAGCCA GGAGGCTAGC AACAGTCTTC ATTCAACCGG CGTCACAATA GTGAGTACCA ATACCGTCGT
1301     CTGCATAATT CTCTTACTGT CATGCCATCC GTAAGATGCT TTTCTGTGAC TGGTGAGTAC TCAACCAAGT CATTCTGAGA ATAGTGTATG CGGCGACCGA
         GACGTATTAA GAGAATGACA GTACGGTAGG CATTCTACGA AAAGACACTG ACCACTCATG AGTTGGTTCA GTAAGACTCT TATCACATAC GCCGCTGGCT
1401     GTTGCTCTTG CCCGGCGTCA ATACGGGATA ATACCGCGCC ACATAGCAGA ACTTTAAAAG TGCTCATCAT TGGAAAACGT TCTTCGGGGC GAAAACTCTC
         CAACGAGAAC GGGCCGCAGT TATGCCCTAT TATGGCGCGG TGTATCGTCT TGAAATTTTC ACGAGTAGTA ACCTTTTGCA AGAAGCCCCG CTTTTGAGAG
1501     AAGGATCTTA CCGCTGTTGA GATCCAGTTC GATGTAACCC ACTCGTGCTC CCAACTGATC TTCAGCATCT TTTACTTTCA CCAGCGTTTC TGGGTGAGCA
         TTCCTAGAAT GGCGACAACT CTAGGTCAAG CTACATTGGG TGAGCACGAG GGTTGACTAG AAGTCGTAGA AAATGAAAGT GGTCGCAAAG ACCCACTCGT
1601     AAAACAGGAA GGCAAAATGC CGCAAAAAAG GGAATAAGGG CGACACGGAA ATGTTGAATA CTCATACTCT TCCTTTTTCA ATATTATTGA AGCATTTATC
         TTTTGTCCTT CCGTTTTACG GCGTTTTTTC CCTTATTCCC GCTGTGCCTT TACAACTTAT GAGTATGAGA AGGAAAAAGT TATAATAACT TCGTAAATAG
1701     AGGGTTATTG TCTCATGAGC GGATACATAT TTGAATGTAT CTAGAAGGTA
         TCCCAATAAC AGAGTACTCG CCTATGTATA AACTTACATA GATCTTCCAT
```

FIGURE 6B CONT.